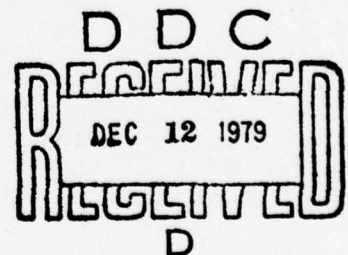Research Memorandum 79-6

LEVEL II

①

FORMATIVE UTILIZATION OF A MODEL FOR THE PREDICTION OF THE EFFECTIVENESS OF TRAINING DEVICES

Marshall A. Narva

EDUCATIONAL TECHNOLOGY AND SIMULATION TECHNICAL AREA

ari

U. S. Army

Research Institute for the Behavioral and Social Sciences

May 1979

.9 22 5 115

14 ARI-RM-79-6

16 Army Project Number
2Q763731A771

Educational Concepts
and Evaluation

9 Research Memorandum 79-6

6 FORMATIVE UTILIZATION OF A MODEL FOR THE PREDICTION
OF THE EFFECTIVENESS OF TRAINING DEVICES.

10 Marshall A. Narva

Leon Nawrocki, Work Unit Leader

12 24

Submitted by:
James D. Baker, Chief
EDUCATIONAL TECHNOLOGY AND SIMULATION TECHNICAL AREA

11 May 1979

Approved by:

Frank J. Harris, Acting Director
Organizations and Systems
Research Laboratory

Joseph Zeidner, Technical Director
U.S. Army Research Institute for
the Behavioral and Social Sciences

Research Memorandums are informal reports on technical research
problems. Limited distribution is made, primarily to personnel en-
gaged in research for the Army Research Institute.
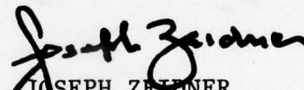
408010

FOREWORD

One of the work units of the System Embedded Training Development project is concerned with the development of a methodology for systematically assessing the characteristics of training devices under development. This work unit is responsive to HRN 78-83, Training Device Concept/Prototype Validation System (TRAINVICE), with the Army Training Support Center as sponsor.

Results of ongoing work are often presented at professional meetings. This paper was presented at the Army Operations Research Symposium at Ft Lee, Virginia on 12 October 1977.

This presentation emphasizes the rationale for development of the methodology. A later presentation of the model, with emphasis on the procedures in its utilization, is given in Research Memorandum 79-7.

JOSEPH ZEIDNER
Technical Director

Formative Utilization of a Model for the Prediction of the Effectiveness
of Training Devices

As a consequence of increasing budgetary and environmental
restrictions associated with the conduct of military operations for
training purposes there has developed an increased interest in the
utilization of training devices, in lieu of or in conjunction with such
operations. However, the development of the devices themselves may well
entail the investment of substantial resources. Therefore, a need
exists for a methodology for the systematic assessment of the character-
istics of training devices under development. If such a methodology
could be developed which would permit making valid predictions or recom-
mendations concerning the selection or design of device concepts or
prototypes, resources could be directed into the development of devices
having the highest probability of leading to the best training results.

An ~~auspicious~~ attempt at ~~such~~ a model, called TRAINVICE, has been
developed for the Army Research Institute (Wheaton, et al., 1976, a, b,
c). This model is based on an extensive review of the literature and is
the result of analytical work by a team of experienced behavioral sci-
entists. This paper will outline the original TRAINVICE model, its
applications, and present suggestions and rationale for a revised model
based upon a formative utilization of TRAINVICE. The revision was
undertaken with a view to enhancing the validity and practicality of
application of the original model, based upon experience gained in its
utilization. Further, the suggested revision aims to make the methodol-
ogy more amenable to utilization by a wider spectrum of users.

The TRAINVICE Model

It would be well to review the TRAINVICE model at this time as it
served as the foundation for the development of the revised model which
will be discussed subsequently. A schematic representation of the
TRAINVICE model is presented in Figure 1. This model is based on the
assumption that certain attributes to be assessed in the training sit-
uation will lead to transfer of training to the operational situation.
Therefore, the higher the rating on the assessment factors, the higher
the transfer that will take place and the more effective the device.
The model provides a framework for the making of these judgments con-
cerning these attributes and combines the results of the judgments. The
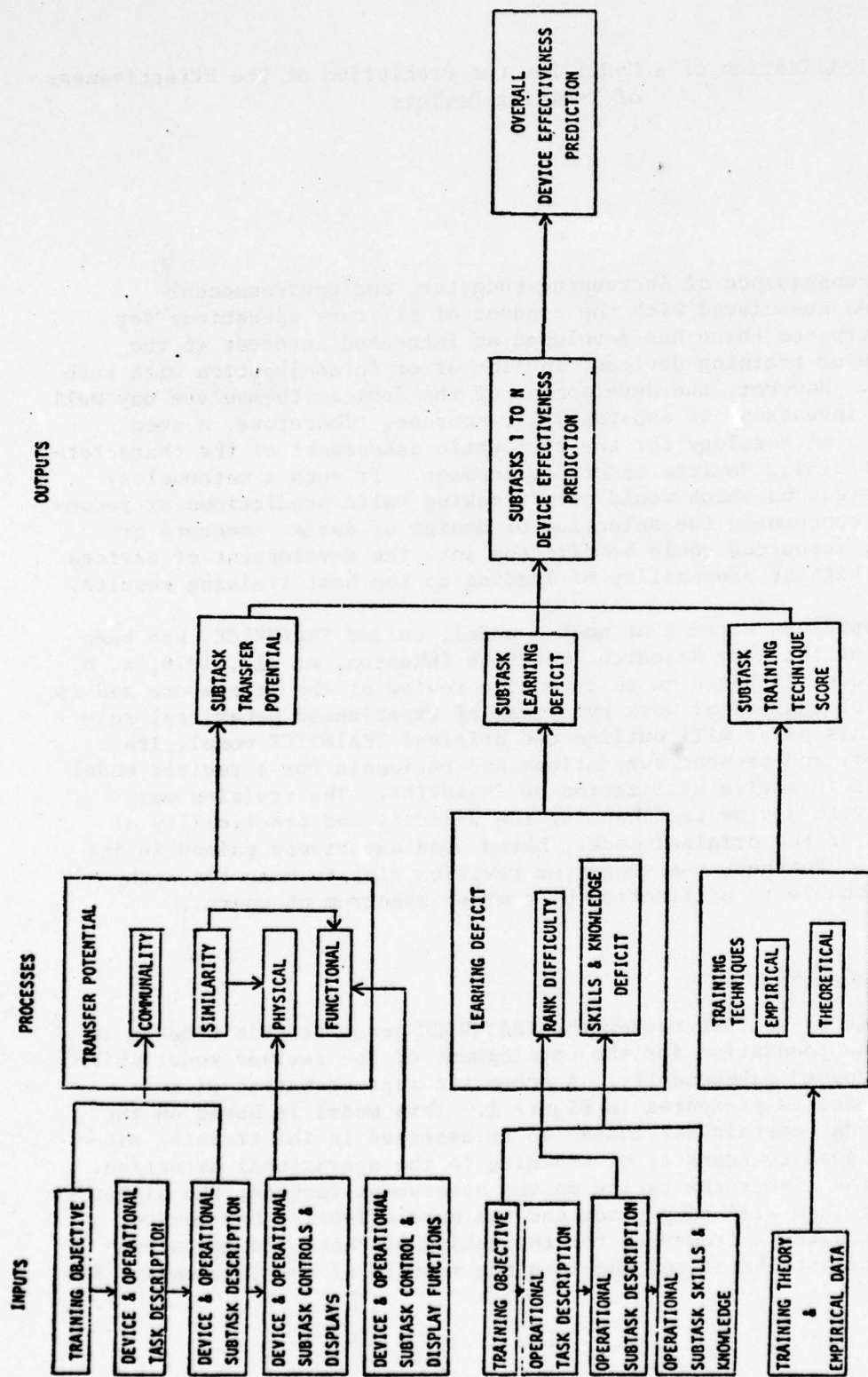
Figure 1. The TRAINVICE Model
(From Wheaton, et al, 1976 b)

2

three variables entering into the assessment are (1) the transfer potential of the device, (2) the learning deficit to be overcome and (3) instructional effectiveness. As with any model, its effectiveness depends on the adequacy of the input data. Inputs into the model consist of descriptions of the tasks and subtasks represented in the operational situation, as circumscribed by the training objective, and those represented in the training device. The controls and displays and their functions for both situations are listed. In addition the skills and knowledge involved in each subtask in the operational situation are formulated for use in the model. Using these inputs, judgments are made using the rating scales given in Figure 2. The subtasks in the two situations, operational and training, are compared to ascertain if provision is made for representation of the subtasks in the training device, in the communality analysis. Next, the displays and the controls for both situations are compared, on physical and functional similarity. As may be seen from the rating scales, the more similar the display or control in the training device is to the operational situation, the higher the score. This is based on the premise that the greater the physical or functional similarity, the greater the transfer of training that will result. Physical similarity refers to the appearance and physical aspects of the displays and controls involved; i.e., their "fidelity"; functional similarity refers to the amount of information conveyed by the display or involved in the operation of a control, in information processing terms. The learning deficit analysis is based upon (1) the assessment of the level of proficiency in each skill or knowledge for the students upon entering the training situation, (2) the desired level of proficiency in each skill or knowledge for the students upon leaving the training situation, and (3) the difficulty (in terms of training time) of training in the skills or knowledges involved in a subtask. This analysis yields a weighted learning deficit for each subtask. The judgments concerning the level of each skill or knowledge are made using the scales shown in Figure 2, adapted from Demaree (1961). The last analysis involved in the TRAINVICE model is an assessment of how well the training device adheres to "good" training techniques. In order to perform this analysis, each of the subtasks is cast into one or more categories of behavior. As given in Figure 2, these categories are those of Braby et al (1972), which are derived from an earlier behavioral categorization by Willis and Peterson (1961). For each of the behavioral categories represented in the subtask, a list of guidelines, also those of Braby et al (1972), are consulted and judgments made, using the scale shown in Figure 2, of the degree to which the guidelines are followed, or not followed, relative to the manner in which the subtask is represented in the training device. The guidelines are broken up into those dealing with the stimulus, response, and feedback aspects of the training situation. For each subtask, the lowest obtained score on each of the three aspects is used to derive an average training technique score. All of the preceding ratings, after conversion to yield a score ranging from 0 to 1, are then fed into an equation to formulate an index of prediction of training effectiveness, ranging from 0 to 1. This equation is as follows:

3

TASK COMMONALITY
1 Present
0 Absent

PHYSICAL SIMILARITY
3 Identical
2 Similar
1 Dissimilar
0 Missing

FUNCTIONAL SIMILARITY
3 Identical
2 Similar
1 Dissimilar
0 Missing

LEARNING DEFICIT ANALYSIS

REPERTORY SCALE
0 No experience, training, familiarity, etc.
1 Has only a limited knowledge of subject or skill
2 Has received complete briefing on subject or skill
3 Understands subject or skill to be performed
4 Has complete understanding of subject or skill

CRITERION SCALE
0 Should have no experience or training
1 Should have limited knowledge of subject or skill
2 Should have received complete briefing on subject or task
3 Should have understanding of subject or skill to be performed
4 Should have complete understanding of subject, or be highly skilled

TRAINING TECHNIQUE ANALYSIS

3 Optimal implementation of this technique
2 Good implementation of this technique
1 Fair implementation of this technique
0 Principle inapplicable or device neither implemented nor violated this principle
-1 Mild violation of this training principle
-2 Serious violation of this principle
-3 Complete violation of this principle

Behavioral Categories:
Recalling facts & principles
Recalling procedures
Nonverbal identification
Nonverbal detection
Using principles, interpreting
Making decisions
Continuous movement
Verbal detection, identification
Positioning & Serial movement
Repetitive movement
Written verbalization
Oral verbalization
Other verbalization

Figure 2. Rating Scales Used in TRAINVICE Model.

4

$$\frac{\Sigma(Ci \times Si \times Ti \times Di)}{\Sigma\ Di}$$

where C is task commonality, S similiarity, T training techniques, and D the training deficit scores for each subtask. The equation was derived from a transfer of training equation of Gagne, Foster and Crowley (1943), which was for use with empirical data, while the TRAINVICE extrapolation deals with judgments made concerning aspects of a device assumed to bring about subsequent transfer of training.

A validation study has been performed on the model, utilizing data obtained during the course of two field studies as criteria against which to compare the predictions derived from use of the model (Wheaton, et al., 1976). The devices were tank gunnery trainers involved with burst-on-target techniques and tracking with the main gun of the M60A1 tank. In each case, the prediction of no differences between the training devices involved was found to be consistent with the equivalence in transfer actually found in utilization of the various devices. This was felt to be a promising but not definitive finding.

In order to obtain additional validation data on the model, and also to obtain experience in utilization of the model to determine if there were aspects that might be changed in order to enhance the practicality of utilization of the model, three Army Research Institute personnel applied the model to two maintenance trainers undergoing evaluation at the Army Ordnance Center and School.[1] This afforded the opportunity to obtain data within a different context than that dealt with by gunnery trainers. These trainers were concerned with automotive troubleshooting. No difference in training effectiveness was predicted for the two trainers, which agreed with the results of the empirical evaluation. Various aspects of the model which caused difficulty in its utilization were noted and influenced the development of the modified version. In addition, ARI conducted a three-day workshop, in which the developers of the original model and individuals who had utilized the model or had an interest in its utilization participated, and this furnished further ideas for possible modification.[2]

The Revised Model

A schematic representation of the main components of the revised model is given in Figure 3.[3]

---

[1]The assistance of Mr. James Dees of the Ordnance Center and School for arranging access to the trainer evaluation is gratefully acknowledged.

[2]The revised model does not necessarily reflect the views of the participants in the workshop. Proceedings of the workshop are in preparation as an ARI report.

[3]The author would like to express special appreciation for the technical review and inputs provided by Ms Barbara Mroczkowski, TRADOC/TSC.

Figure 3. The Revised Model

Assign Skills/Knowledge to Behavioral Categories

Generic Characteristics Listing

Physical Characteristics Guidelines

Physical Characteristics Analysis

Functional Characteristics Guidelines

Functional Characteristics Analysis

Derive Device Effectiveness Index

Coverage Requirement Analysis

Coverage Analysis

Training Criticality Analysis

Training Difficulty Analysis

Training Objective

Operational & Device Task Description

Operational & Device Subtask Description

Operational & Device Subtask Skills & Knowledge

Operational & Device Subtask Cues & Responses

Consolidate Operational & Device Skills & Knowledge
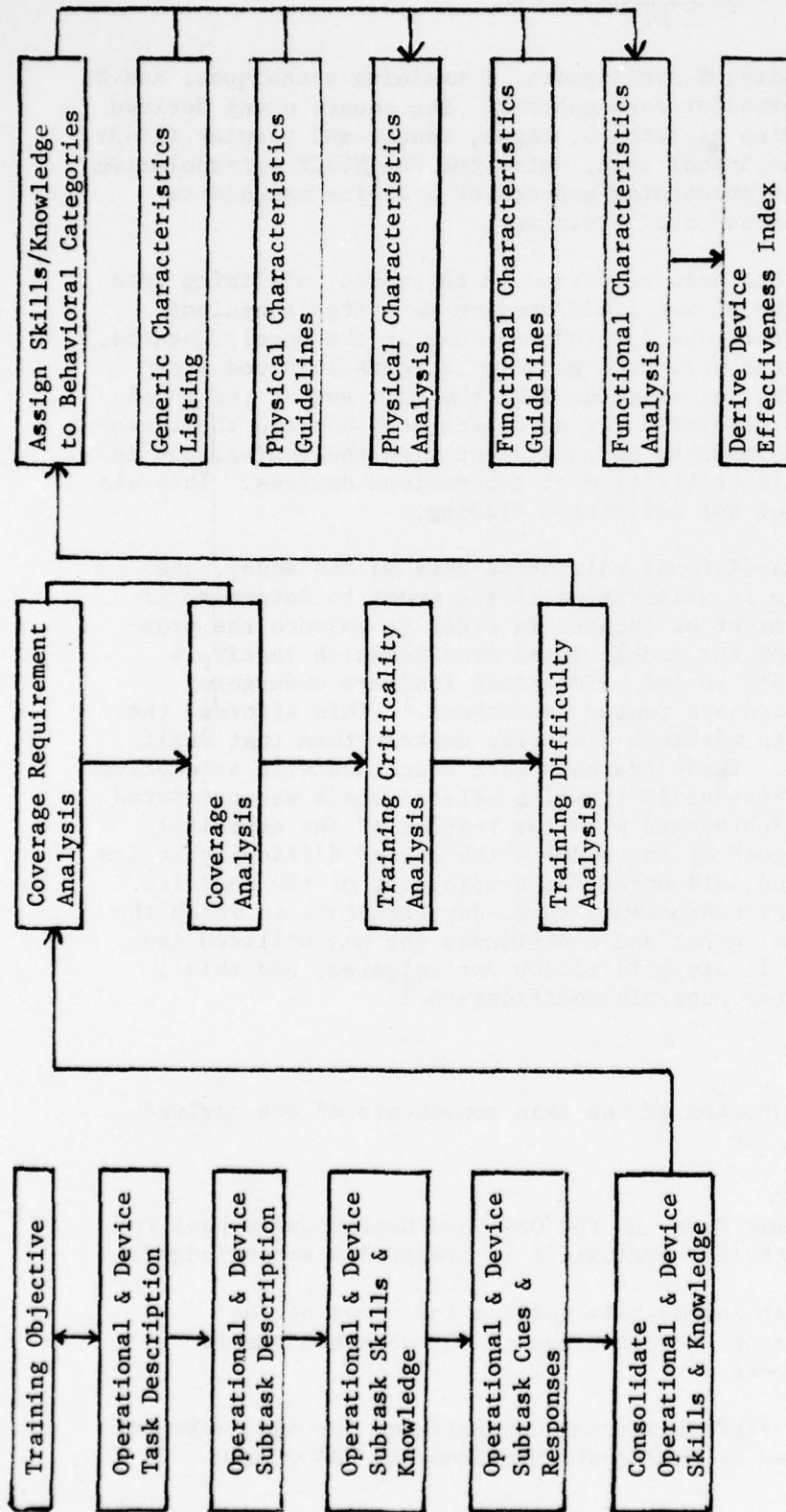
6

Basically, the model considers three main aspects of the training device during the course of the assessment; what, why, and how. "What" involves an analysis of what skills (or knowledge) should be covered in the training situation and what skills are covered, in order to ascertain that the spectrum of skills covered neither exceeds that which is necessary nor leaves out any that should be covered. It determines that the skills covered are in keeping with the training objective. The "why" refers to that aspect which determines why these skills should be covered, apart from their being included within the spectrum of skills subsumed under the training objective. This consists of two main aspects; training criticality, which relates to the degree of proficiency required in each skill at the end of the training, and training difficulty, which considers the degree of difficulty to be expected in training for each skill. These two analyses, in essence, give a weight to each skill. The last main analysis, refers to the "how" of the training device; that is, how is each skill taught, and compares the training situation to guidelines of "good" practice. These guidelines are applied to two aspects of the training device; the physical and the functional.

As with the TRAINVICE model, the assumption is being made that the potential for transfer of training will increase as a function of the degree to which the skills are represented, within the constraints of the training objective, and the degree to which the training situation follows guidelines for "good" practice. In addition each of the skills is appropriately weighted by degree of skill needed and degree of difficulty.

Input data requirements. The first requirement is the statement of the training objective. As presented in TRADOC Reg 350-100, Systems Engineering of Training (1973), or in the Interservice Procedures for Instructional Systems Development (1975), the training objective states the action, or task, that the student should be capable of performing, and the conditions and standards of performance he is to attain as a result of the training. The training objective carves out a piece of the operational situation that is to be subjected to training and determines to what level this segment is to be trained, which may vary with relationship to the level actually required in the operational situation. This training objective may vary as to level of specificity, with the subsumed task and associated subtasks also varying as to specificity. It should be kept in mind that the procedures given in such publications as TRADOC Reg 350-100-1 or TRADOC Pam 350-30 (ISD model) deal with the derivation of a program of instruction, of which a training device may be one small segment. The training device may be considered to be but one of many possible media or instructional delivery systems. Therefore, the procedures must be utilized at a level of specificity suitable for the assessment of a training device. Guidance to the conduct of a task analysis are given in such publications as those above and those of R. Miller (1953a), Rankin (1975) and Chenzoff and Folley (1965). However, no amount of guidance can substitute for the good judgment of the analyst in formulating meaningful segments of activity. Information must be

derived concerning the task, component subtasks, required skills and knowledge, and the cues and responses involved in the execution of each of the skills, for both the operational and the training situations. Therefore, the same input requirements exist as for the original TRAINVICE model, with the additional requirement for the definition of skills and knowledge being exercised in the training situation. In TRAINVICE, it may be recalled, the subtasks are broken out for both the operational and the training situations, but the skills and knowledges are broken out only for the operational situation, in conjunction with the derivation of the learning deficit for each of the skills. This was concerned only with the characteristics of the student and did not consider what is offered by the device itself. In the revised model, the unit of concern is at the skills/knowledge level rather than the subtask, although these two units of activity can be very similar. The rationale for going with skills and knowledges as the unit of concern is based on the assumption that the prime objective of a training device is to provide for the acquisition and practice of those skills and knowledge required to carry out the task subsumed in the training objective.

Within each skill or knowledge, the cue(s) and response(s) involved are to be extracted rather than the displays and controls, as in the TRAINVICE model, to provide for greater flexibility in applying the model.

Since the same skills or knowledge may occur over the course of more than one subtask, the skills and knowledges for the operational situation are consolidated into one list in order to avoid duplication. The same is done for the skills and knowledges from the training situation.

Coverage requirements analysis. The first analysis to be performed is that concerned with the requirement for training for each of the skills subsumed under the training objective, from the point of view of the operational situation. The judgment is made for each skill (or knowledge) as to whether it should be included or not be included in the training situation. Depending upon the complexity of the training objective, this judgment may or may not be straightforward. With a fairly constrained or simple training objective or task, most of the skills may be readily judged to be necessary for the training situation. In some cases, it may be necessary to postpone this analysis until the subsequent analyses dealing with training criticality and training difficulty have been made; in some cases judgment concerning training coverage requirement may be modified as a result of the subsequent criticality and difficulty judgments. It should be noted that this analysis does not deal with the mission criticality of the subtask, it is assumed that all the subtasks are necessary to mission success, but rather with the necessity for providing training for the skills subsumed within each subtask. This analysis is a "gate" only; it determines if the skill should be represented in the training. Depending upon the stage of development of the training device, this analysis may help to delineate the range of skills to be represented in a device as well as

8

assessing the range of skills represented in a device. It may also help
to define or modify the training objective, if the stage of development
of the training program permits. If the skill is judged to require its
presence in the training situation, it is given a "1"; if not it is
given a "0". If, due to the stage of development, this analysis is not
considered necessary, all skills would be rated "1." It should be noted
that this judgment is similar to that made during the course of the
learning deficit analysis in the TRAINVICE model. In assessing the
level of the skill required at the close of training, provision was made
for a "0" rating, which indicated that no training was required in the
skill under consideration; it was the low end of the scale. This type
of judgment has been broken out into this separate coverage requirement
analysis in the revised model to permit a more complete assessment of
the range of skills in the training situation.

Coverage Analysis. In conjunction with the Coverage Requirement
analysis, a Coverage Analysis is also performed. Comparing the opera-
tional and training lists of skills, an assessment is also made for each
skill in the operational situation as to whether it is represented in
the training situation. If it is, it is given a rating of "1," if it is
no, it is given a "0." This analysis is the same as the Communality
analysis performed in the TRAINVICE model. As in TRAINVICE, both actual
and potential coverage may be determined depending upon how the device
is utilized. However, this analysis is made concerning skills rather
than subtasks, and the list of operational skills has been "adjusted" by
the coverage requirement analysis performed previously.

If the coverage requirement rating is "1" and the coverage rating
is "0," this indicates that training in this skill is lacking and steps
should be taken to include it, if possible, or the device will suffer in
the overall rating. On the other hand, if the coverage requirement
rating is "0" and the coverage rating is "1," this would indicate that
unnecessary training is being provided and should be eliminated from the
device, if possible, or the overall rating of the device will suffer.

If a "0" rating is given for a skill either in the Coverage Requirement
or Coverage Analysis, no further analysis for that skill need be done as
the overall rating for that skill reduces to "0" due to the multiplica-
tive nature of the derived index.

Training Criticality Analysis. For each of the skills that have
earned a "1" rating on both the Coverage Requirement and the Coverage
Analyses; that is, for those skills that have been judged to be necessary
in the training situation and are indeed represented, a judgment is made
as to the degree of proficiency required in that skill at the end of
training. Ratings are made using the scale shown in Figure 4. This is
the same scale as utilized in the TRAINVICE model for the criterion
scale used in the Training Deficit Analysis, with the exception of the
"0" point on the scale.

COVERAGE REQUIREMENT
0 Not required
1 Required

TRAINING CRITICALITY
1 Should have limited knowledge of subject or skill
2 Should have received complete briefing on subject or skill
3 Should have understanding of subject or skill to be performed
4 Should have complete understanding of subject, or be highly skilled

Behavioral categories:
Rule learning and using
Classifying-Recognizing Patterns
Identifying symbols
Detecting
Making decisions
Recalling bodies of knowledge
Performing gross motor skills
Steering & guiding-continuous movement
Positioning movement & recalling procedures
Voice communicating

COVERAGE
0 Not covered
1 Covered

TRAINING DIFFICULTY
1 Minimal or none
2 Some
3 Much
4 Substantial

DEVICE CHARACTERISTICS ANALYSIS

PHYSICAL CHARACTERISTICS
0 Not adequate for requirements/guidelines
1 Adequate implementation for requirements/guidelines
2 Good implementation for requirements/guidelines
3 Outstanding implementation for requirements/guidelines

FUNCTIONAL CHARACTERISTICS
0 Not adequate implementation of guideline
1 Adequate implementation of guideline
2 Good implementation of guideline
3 Outstanding implementation of guideline

Figure 4. Rating Scales Used in Revised Model

As noted previously, this scale is adapted from Demaree (1961). Once again, it is to be noted that this analysis is concerned with training criticality and not mission criticality, although mission criticality may enter into the determination of training criticality. As noted in Reg 350-100-1, the standards for skills are usually derived from the standard established for the task or subtask. However, this standard must be tempered by the nature of the skill and the degree to which the training on the skill will be merged with subsequent on-the-job training and the degree to which the training is embedded in other aspects of the training program than that involving exposure to the device.

Training Difficulty Analysis. In addition to assessing the level of proficiency required for each of the required skills, an assessment is made, in the Training Difficulty Analysis, of the degree of difficulty which is to be expected in attaining that level of proficiency, for the particular skill and trainee population involved. As seen in Figure 4, a scale with four points is used. This scale is modified from that of Rankin (1975). In determining this rating, the following factors must be taken into account, the inherent difficulty of the skill, the amount of proficiency in the skill by the trainee population due to prior training or experience, and the level of proficiency required at the end of the training, as reflected in the training criticality rating. While no formal procedures are recommended for the assessment of each of these factors, each must be taken into account by the analyst. This is a similar procedure to that followed in deriving the weighted learning deficit score in the TRAINVICE model. It may be recalled that for that derivation, a rating is given for each skill relative to the degree of proficiency held by the trainees at the beginning of the training and also at the end of the training and the two are substracted. Then the subtasks are rank ordered on difficulty. The ranks were then multiplied by the mean subtask deficit to give the weighted deficit score. It was felt during the application of the TRAINVICE model to the maintenance trainers that the rank ordering of the subtasks on difficulty was difficult. In the revised model, each of the skills is rated against an absolute scale of difficulty.

Device Characteristics Analyses. Up to this point we have been concerned with assessing what skills are covered by the device, their "fit" to the training objective, and why the skills have to be included in the training situation. We now turn our attention to how these skills are to be taught. In order to do this, attention must now be turned directly to the device and its characteristics. Up to now, the device has been considered only from the point of view of the coverage of the desired skills and knowledge that it offers. It may be recalled that the TRAINVICE model looked at the displays and controls in the Physical and Functional Similarity analyses and at the subtasks, translated into behavioral categories, in terms of how well certain principles were utilized by the device. In essence, the physical and the functional characteristics of the device were considered, and it is these two

11

aspects that are broken out and considered in the Device Characteristics Analyses of the revised model. Therefore there are two analyses; the Physical Characteristics Analysis and the Functional Characteristics Analysis.

    Physical Characteristics Analysis. In the TRAINVICE model, a Physical and a Functional Similarity analysis was performed. It may be recalled that the more similar the physical or functional aspect of the display or control on the training device relative to the corresponding display or control in the operational situation, the higher the rating given. This was based on the premise that the greater the similarity, the greater the transfer potential. However, this may be called into question. As R. Miller (1954, p. 29) has pointed out, "The design of training devices should be directed towards maximum transfer of training value, not physical realism." He further states, "Some stages of training and kinds of task trained require less physical realism than others," and "The kinds and extent of physical realism built into a given training device should be based upon careful examination of what is psychologically important." Demaree (1961, p. 44) has said that "...the general rule is to represent the operational equipment so that realism is attained with regard to what is to be learned and not to the operational equipment." In other words, more specific critera are required for the evaluation of the transfer potential of the physical characteristics of the training device, in terms of the stage of learning and type of behavior involved than a simple correlation with the physical characteristics of the operational equipment. R. Miller (1953b, c), Demaree (1961), G. Miller (1974), Kinkade and Wheaton (1972), Lumsdaine (1960), Muckler (1959), Caro (1976), Klein (1976) and Smode (1971) are among those who have addressed the problem of the physical characteristics or "fidelity" of training devices. They do give recommendations; however, for the most part they consider classes of devices, or gross characteristics, and do not use the same classifications of devices or behaviors. It is difficult to extrapolate to the generic characteristics of devices and to a particular device. Much work is still needed to come up with guidelines to assist the analyst in assessing a specific display or cue presented by a device, or to compare one device within a particular class of devices with another device as to its transfer potential. However, the procedure followed in the original TRAINVICE model of correlating physical and functional similarity with transfer potential was rejected as being too rigid and indeed possibly misleading. While Caro (1970) had advocated a procedure in which the stimuli and responses in the operational and training situations are compared, in which positive transfer was assumed to occur when both stimuli and responses were similar, he was also concerned with the stimulus-response pairing, predicting negative transfer when the stimuli were similar but the responses to the similar stimuli were different, something not adequately considered in the TRAINVICE procedure. Also, the Caro analysis was performed within the context of instrument flight simulators, a complex perceptual-motor behavior that may indeed call for a high degree of realism in the training, but which may not represent a valid requirement for other types of

12

behaviors. In casting about for guidance for a procedure to assess the physical characteristics of a device, it was decided to combine the procedure suggested by Braby, et al (1975) of the Navy Training Analysis and Evaluation Group in conjunction with their Training Effectiveness Cost Effectiveness Prediction (TECEP) technique with selected guidelines from the Interservice Procedures for Instructional Systems Development (the ISD model)(TRADOC 1975). These guidelines, it should be noted, are a simplified version of the learning guidelines developed by Aagard and Braby (1976) in conjunction with the TECEP model. Various sets of guidelines were considered; those formulated by Willis and Peterson (1961), Gagne and Briggs (1974), an earlier set by Braby et al (1972) (which was the set used in the Training Techniques Analysis of the TRAINVICE model), the Aagard and Braby (1976) guidelines, and the simplified version of their guidelines given in the ISD model (TRADOC 1975). In order for the revised model to have widest applicability, the guidelines from the ISD model were adopted. These guidelines deal for the most part with functional aspects of the training situation, such as the sequencing of learning events. However, selected guidelines were chosen as being applicable to the design of specific elements of the device or training situation. These specific elements, be they displays-controls, inputs-outputs, or cue-response pairs may be likened to the simulation elements of Smode (1971). The adequacy of these simulation elements determine the perceptual equivalence of the training and operational environments. They are the elements in the total mosaic of the training device, which essentially is a spatial and temporal placement and sequencing of these elements and, as Matheny (1974) has pointed out, the assumption may be made that it is perceptual equivalence that results in positive transfer. To bring the number of possible specific forms that each element may assume into manageable proportions, it is necessary to translate the specific skills to be represented on the training device into behavioral categories and the specific simulation elements must be translated into generic characteristics.

The guidelines extracted from those in the ISD model give limited guidance to the analyst as to the specific physical characteristics that the cues and responses should take for maximum transfer potential for the type of behavior involved. Therefore, the analyst must make a judgment as to what generic characteristics are required. His judgment as to what is required is merged with his assessment of how well the generic characteristics of the cue or response do follow the available guidelines. In order to make these judgments, a list of generic characteristics dealing with the stimulus and response characteristics of the training device are used. This listing, given in Figure 5, is taken from the listing given in the TECEP technique of Braby, et al (1975). To make this judgment, the skill is first placed into one of the behavioral categories, shown in Figure 4, which was also taken from the TECEP technique, and which is also utilized in the ISD model. This permits access to a list of guidelines to the physical characteristics deemed desirable for each of the behavioral categories. Then the cues and associated responses subsumed under each skill are considered, utilizing the list

13

STIMULUS CAPABILITIES

Visual Form

1. **Visual Alphanumeric** - words, numbers and other symbols presented graphically.

2. **Visual Pictorial, Plane** - a two-dimensional image, a representation in the form of a photograph or drawing.

3. **Visual Line Construction, Plane** - a two-dimensional figure made of lines, such as a mathematical curve or graph.

4. **Visual Object, Solid** - a three-dimensional image or reality that is viewed from exterior perspectives.

5. **Visual Environment** - A three-dimensional image or reality that is viewed from inside.

Visual Movement

6. **Visual Still** - a static visual field, as with a still photograph, drawing, or printed page.

7. **Visual Limited Movement** - a basically static visual field with elements that can be made to move, as with an animated transparency or simple panel with switches that move.

8. **Visual Full Movement** - a visual field in which all elements can move, as with a motion picture, flight simulator, or operational aircraft.

9. **Visual Cyclic Movement** - a visual field which moves through a fixed sequence and then repeats the sequence in a repetitive manner, as with a film loop.

Visual Spectrum

10. **Black and White** - a visual field composed of either black or white elements, as with the printed page or line drawings.

11. **Gray Scale** - a visual field composed of black, white and continuous gradations of gray, as with a black and white photograph or television picture.

12. **Color** - a visual field composed of various segments of the visual spectrum, as with color television or motion pictures.

Scale

13. **Exact Scale** - actual visual field or a one-to-one replication of that field as with a full-sized mock-up, simulator, or operational system.

14. **Proportional Scale** - a representation of reality in other than full scale, such as a scaled model map or photograph.

Audio

15. **Voice Sound Range** - a limited quality of sound which enables spoken words to be used as the medium of communications, but not suited to more demanding tasks, such as music or sound recognition exercises.

16. **Full Sound Range** - a quality of sound reproduction that contains all the significant elements of the sound and is suited to the demanding task of sound recognition exercises.

17. **Ambient Sounds** - a complex sound environment with sounds emanating from various sources and from various directions, including background noise and task significant sounds.

Other

18. **Tactile Cues** - signals received through the sense of touch, including sensations related to texture, size or shape.

19. **Internal Stimulus Motion Cues** - the sensations felt by a person when he moves his arm, leg, fingers, etc.

20. **External Stimulus Motion Cues** - the sensations felt by a person when he is moved by some outside force in such a way that his body experiences roll, pitch, yaw, heave, sway and/or surge.

TRAINEE RESPONSE MODES

21. **Covert Response** - a response which the trainee creates in his mind but does not express in an observable manner.

22. **Multiple Choice** - a response mode in which a trainee selects a response from a limited set of responses.

23. **Pre-programmed Verbal Performance** - a response mode in which a trainee creates a short answer to a question having a limited set of correct answers.

24. **Free-Style Written Performance** - a response mode in which a trainee writes a response in his own words.

25. **Decision Indicator** - a verbal or perceptual motor response in which the trainee indicates that he has made a divergent type decision.

26. **Voice Performance** - a response mode in which a trainee speaks, including conversation.

27. **Fine Movement Manipulative Acts** - a response mode in which a trainee makes discrete and small movements of dials, switches, keys or makes sensitive adjustments to instruments. Act may involve use of small instruments.

28. **Broad Movement Manipulative Acts** - a response mode in which a trainee makes large movements of levers or wheels on large pieces of equipment or by the use of hand held tools.

29. **Tracking** - a response mode in which a trainee continuously controls a constantly changing system, such as steering an automobile or holding a compass bearing in steering a ship.

30. **Procedural Manipulative Acts** - a response mode in which a trainee performs the sequence of steps in a procedure, such as in the carrying out of the items on the checklist for pre-flighting an aircraft or turning on a radar system.

Figure 5. Generic Characteristics List Used in Revised Model
(From Braby et al, 1975)

14

of generic characteristics. The cue or response is categorized into
each of the applicable generic characteristics and to each characteris-
tic a rating is given, using the scale shown in Figure 4. This scale
ranges from a "0" rating which means that that generic characteristic is
not adequate from the point of view of the analyst's judgment of what is
required and/or the implementation of the applicable guideline to a
rating of "3" which represents an outstanding implementation. Therefore,
each cue-response pair involved in each skill represented by the train-
ing device receives a cluster of ratings on the applicable generic
characteristics. The pattern of these ratings will serve to "highlight"
the various physical characteristics of the device, both the outstanding
aspects and those that require change. They may serve as a profile of
the characteristics, much as that proposed by Smode (1971). In order to
derive the Physical Characteristics rating for the skill involved, the
ratings given on each of the generic characteristics are added to give
the total for that skill. Therefore, the presence of a "0" rating does
not eliminate that skill from the total rating but does serve to down-
grade the total rating for the skill. In order to derive a baseline
against which the rating may be compared in the final device rating, the
number of generic characteristics involved is multiplied by "3," the
highest possible rating, to give the maximum possible rating for that
skill as far as Physical Characteristics of the device are concerned.

    The requirement imposed upon the analyst to make a judgment as to
the generic characteristic requirements follows that utilized in the
TECEP technique. That technique, however, is concerned with the selec-
tion of an instructional delivery system, of which a training device is
but one alternative. Indeed, the TECEP technique gives recommendations
as to which delivery systems permit the application of the learning
guidelines for each of the behavioral categories; recommendations based
on the pattern of matching of the generic characteristics inherent in
the various delivery systems and those judged to be necessary by the
analyst. Jorgensen (1976) has utilized a similar matrix approach in
which the required generic characteristics are matched against various
media, of which training devices are but one class, in order to select
training concepts (or media) most suitable for training various tasks.
Similar procedures may be found in the ISD model (TRADOC, 1975) (which
is taken from the TECEP technique), the Air Force Handbook for Designers
of Instructional Systems (1974), Parker and Downs (1961), Nunnelly
(1966), Bennik and Hoyt (1977), Bretz (1971) and the review by Spangenberg,
Riback and Moon (1973). However, these procedures are intended for the
selection of or comparison of various media or instructional delivery
systems rather than the scrutiny of a training device per se or compari-
son of training devices.

    It will be recalled that in the TRAINVICE model, the analyst looks
at the displays and controls and compares each with its counterpart in
the operational situation and gives it a rating on physical and functional
similarity. The physical similarity rating has been replaced in the
revised model by the procedure given above for the reasons discussed

15

above. The functional similarity analysis has been dropped as it was found in the application to the maintenance trainers to pose a difficult decision, not suitable for a procedure aimed at a wide spectrum of users, and was essentially tied to the physical characteristic, specifically being determined by the number of states (which in many cases were difficult to determine) which could be assumed by the display or control.

Functional Characteristics Analysis. The other analysis subsumed under the Device Characteristics Analysis is that concerned with the functional characteristics. While the Physical Characteristics analysis is concerned with the analysis of the elements of the training device per se the functional analysis is concerned with how these elements are utilized. This analysis may be compared to the Training Techniques Analysis of the TRAINVICE model. As in that analysis, the operations of the device are compared against guidelines to ascertain to what extent "good" training practices are followed. Certain changes have been made, however. Instead of using the subtask as the unit of concern, each skill is analyzed. It was found during the application of the model to the maintenance trainer, that there often was confusion and at times even conflicting guidelines to be considered, as the subtask was broken up into several behavioral categories, each with its own set of guidelines. In the revised model, each skill is translated into one of the behavioral categories (those used in the ISD model), ss shown in Figure 4, and only that set of guidelines is considered. The guidelines have been changed to those used in the ISD model, as opposed to the earlier Braby guidelines which were used by the TRAINVICE model. These guidelines are felt to be more straightforward and more suitable for a wider spectrum of users.

Therefore, to perform this analysis, each of the skills is translated into one of the behavioral categories given in Figure 4. The appropriate set of guidelines is consulted, as the functional, dynamic characteristics of the elements involved in training for that skill are considered. First a determination is made if the particular guideline is applicable. If it is, a rating, using the scale shown in Figure 4, is made. As in the Physical Characteristics analysis, the scale ranges from "0" which means that the guideline is not adequately followed in the training for that skill up to a rating of "3" which represents a judgment of outstanding implementation of the guideline. Therefore, each of the applicable guidelines receive a rating for that skill. In order to derive the Functional Characteristics rating for the skill, the rating given for each of the applicable guidelines are added to give the total for that skill. The guidelines not deemed applicable are not considered in the ratings. As the total is derived through addition rather than multiplication of the individual ratings, the presence of a "0" rating does not eliminate that skill from the total device rating, but does serve to downgrade the rating for the skill and also may serve as a "flag" for something that needs to be corrected. It will be recalled that in the TRAINVICE model the lowest rating was used and all other ratings discarded. It was felt that this was a waste of valuable information, and possibly misleading, as the one low rating could obscure the presence of

16

other high ratings. The revised procedure takes all the ratings into
account. In order to derive a baseline against which the rating may be
compared in the final device rating, the number of applicable guidelines
is multiplied by "3," the highest possible rating, to give the maximum
rating for that skill for functional characteristics. This would represent
a device with outstanding application of all the applicable guidelines
for the behavior within which the skill falls.

In applying guidelines during the course of the Training Techniques
Analysis in the use of the TRAINVICE model, some difficulty was encountered
as the unit of activity emcompassed in a subtask did not lend itself to
application of the guidelines, as many guidelines were concerned with a
sequencing of events which covered a more extensive period of time than
covered by the subtask. While this problem may be accentuated by the
use of skills as the unit of activity, the possibility also exists that
since skills exist over subtasks, the match of unit of activity and
guidelines may be enhanced. It depends upon the particular task and
subtasks involved. This problem tends to arise from the fact that the
guidelines, both those used in the original TRAINVICE and the revised
models, originate from those intended to be used for instructional
system development and instructional delivery system or media selection,
of which a training device may be one small segment. The development of
more specific guidelines is required.

Derivation of Index of Predicted Training Device Effectiveness. It
will be recalled that the derivation of the index of the effectiveness
of the training device utilized in the TRAINVICE model was based on one
of the formulas discussed by Gagne, Foster, and Crowley (1948) to express
transfer of training in terms of empirical data. The equation formulated
for use with the revised model also follows a procedure discussed by
Gagne, Foster, and Crowley (1948). While not based on one of their
formulas directly, it is in keeping with their conclusion that the most
useful and practical type of formulation is that based on percentage of
maximum possible transfer. It assumes that if the device were to follow
perfectly all of the guidelines, as judged necessary by the analyst,
that maximum transfer, which could be attributed to the device, would be
the result. This forms the baseline against which the device under
evaluation is compared. Therefore, the maximum possible score for the
Physical Characteristics Analysis and the maximum possible score for the
Functional Characteristics Analysis added together would represent the
"perfect" device for the training of that particular skill. This total
is weighted by the Coverage Requirement, Coverage, Training Criticality
and Training Difficulty scores derived for that skill. The derived
score for each skill is then compared with the score representing maximum
expected transfer. (If a "0" rating is given for either the Coverage
Requirement or Coverage Analysis, the total score for that skill is
reduced to "0" and makes no contribution to the derived index for the
device.) To derive the score for the total device, each of the skill
scores is added. Therefore, the index of predicted training device
effectiveness is as follows:

17

$$\frac{\Sigma(CR \times C \times Ci \times D \times (PC + FC))i}{\Sigma(CR \times C \times Ci \times D \times (PC_{max} + FC_{max}))i}$$

where: CR   Coverage Requirement Score
            C   Coverage Score
            Ci   Training Criticality Score
            D   Training Difficulty Score
            PC   Physical Characteristics Score
            FC   Functional Characteristics Score
            PC_{max} Maximum Possible Physical Characteristics Score
            FC_{max} Maximum Possible Functional Characteristics Score
                     for each skill.

This equation will yield an index ranging from 0 to 1. The larger the index, the higher were the ratings given on the Device Characteristics Analyses and presumably the greater the potential for transfer of training.[1]

_____

[1]This equation was subsequently changed to:

$$\frac{\Sigma(CR \times C \times Ci \times D \times (PC + FC))i}{\Sigma(CR \times C \times 4 \times 4 \times (PC_{max} + FC_{max}))i}$$

18

References

Aagard, A. and Braby, R. Learning Guidelines and algorithms for types of training objectives. TAEG Report No. 23, U.S. Navy, Training Analysis and Evaluation Group, Orlando, FL, March 1976.

Bennik, F. D. and Hoyt, W. G. Determining TEC media alternatives for field artillery individual-collective training in the FY 78-83 period. Task 7 report, TM-5841/001/00, System Development Corp., Santa Monica, CA, 7 June 1977.

Braby, R., Micheli, G., Morris, C., and Okraski, H. Staff study on cost and training effectiveness of proposed training systems. TAEG Report No. 1, U.S. Navy, Training Analysis and Evaluation Group, Orlando, FL, 1972.

Braby, R., Henry, J., Parrish, W., and Swope, W. A technique for choosing cost-effective instructional delivery systems. TAEG Report No. 16, U.S. Navy, Training Analysis and Evaluation Group, Orlando, FL, April 1975.

Bretz, R. The selection of appropriate communication media for instruction: A guide for designers of Air Force Technical Training Programs. R-601-PR, Rand Corporation, Santa Monica, CA, February 1971.

Caro, P. Equipment-device task commonality analysis and transfer of training. Technical Report 70-7, Human Resources Research Organization, Alexandria, VA, June 1970.

Caro, P. The role of simulators in pilot training: some questions regarding effectiveness. Proceedings, Third Triennial Symposium on Cost Effective Learning Through the Application of Technology. Society for Applied Learning Technology, Washington, DC, July 1976.

Chenzoff, A. and Folley, J. Guidelines for training situation analysis (TSA). Technical Report NAVTRADEVCEN 1218-4, U.S. Naval Training Device Center, Port Washington, NY, July 1965.

Demaree, R. Development of training equipment planning information. ASD Technical Report 61-533, U.S. Air Force, Aerospace Medical Laboratory, Wright-Patterson Air Force Base, OH, October 1961.

Gagne, R. and Briggs, L. Principles of instructional design. New York: Holt, Rinehart and Winston, Inc., 1974.

Gagne, R., Foster, H., and Crowley, M. The measurement of transfer of training. Psychol. Bulletin, 1948, 45, 97-130.

Jorgensen, C. A methodology and analysis for cost-effective training in the AN/TSQ-73 Missile Minder. Army Research Institute for the Behavioral and Social Sciences, Ft Bliss Field Unit, El Paso, TX, Research Memorandum 77-26, February 1978.

Kinkade, R. and Wheaton, G. Training device design. In Van Cott, H. and Kinkade, R. (Eds) Human engineering guide to equipment design (Revised), Joint Army-Navy-Air Force Steering Committee, U.S. Government Printing Office, Washington, DC, 1972.

Klein, G. Problems and opportunities in deriving training requirements for the design and utilization of simulators. Proceedings, First International Learning Technology Congress and Exposition, Volume IV, Society for Applied Learning Technology, Washington, DC, July 1976.

Lumsdaine, A. Design of training aids and devices. In Folley, J. (Ed.), Human factors methods for system design. AIR-290-60-FR-225, American Institute for Research, 1960.

Matheny, G. Training simulator characteristics: Research problems, methods, and performance measurements. Proceedings, Aircrew Performance in Army Aviation, U.S. Army, Office of the Chief of Research, Development and Acquisition, July 1974.

Miller, G. Some considerations in the design and utilization of simulators for technical training. U.S. Air Force, Human Resources Laboratory, Brooks Air Force Base, TX, August 1974.

Miller, R. A method for man-machine task analysis. WADC Report 53-137, U.S. Air Force, Aero Medical Laboratory, Wright-Patterson Air Force Base, OH, June 1953a.

Miller, R. A method for determining human engineering design requirements for training equipment. WADC Technical Report 53-135, U.S. Air Force, Aero Medical Laboratory, Wright-Patterson Air Force Base, OH, June 1953b.

Miller, R. Handbook on training and training equipment design. WADC Technical Report 53-136, U.S. Air Force, Aero Medical Laboratory, Wright-Patterson Air Force Base, OH, June 1953c.

Miller, R. Psychological considerations in the design of training equipment. WADC Technical Report 54-563, Aero Medical Laboratory, Wright-Patterson Air Force Base, OH, December 1954.

Miller, R. Task and part-task trainers and training. WADC Technical Report 60-469, Aerospace Medical Division, Wright-Patterson Air Force Base OH, June 1960.

Muckler, F., Nygaard, J., O'Kelly, L., and Williams, A. Psychological variables in the design of flight simulators for training. WADC Technical Report 56-369, Aero Medical Laboratory, Wright-Patterson Air Force Base, OH, January 1959.

Nunnelly, C., Klemmer, A., Corrigan, R., and Kaufman, R. The instructional system approach to maintenance technical training: Development and implementation model. Human Factors, 1966, 8, 163-172.

Parker, J. and Downs, J. Selection of training media. ASD Technical Report 61-473, Aerospace Medical Division, Wright-Patterson Air Force Base, OH, September 1961.

Rankin, W. Task description and analysis for training system design. Technical memo, 74-2, U.S. Navy, Training Analysis and Evaluation Group, Orlando, FL, January 1975.

Smode, A. Human factors inputs to the training device design process. Technical Report NAVTRADEVCEN 69-C-0298-1, Naval Training Device Center, Orlando, FL, September 1971.

Spangenberg, R., Riback, Y., and Moon, H. The state of knowledge pertaining to selection of cost-effective training methods and media. HumRRO Technical Report 73-13, Human Resources Research Organization, Alexandria, VA, June 1973.

U.S. Army, Systems engineering of training (Course design), TRADOC Reg 350-100-1, Headquarters, U.S. Army Training and Doctrine Command, Fort Monroe, VA, 6 July 1973.

U.S. Army, Interservice procedures for instructional systems development. TRADOC Pam 350-30, U.S. Army Training and Doctrine Command, Fort Monroe, VA, August 1975.

U.S. Air Force, Handbook for designers of instructional systems, AFP 50-58, Headquarters, U.S. Air Force, Washington, DC, 1973-1974.

Wheaton, G., Rose, A., Fingerman, P., Korotkin, A., and Holding, D. Evaluation of the effectiveness of training devices: Literature review and preliminary model. Research Memorandum 76-6, U.S. Army Research Institute for the Behavioral and Social Sciences, April 1976a.

Wheaton, G., Fingerman, P., Rose, A., and Leonard, R. Evaluation of the effectiveness of training devices: Elaboration and application of the predictive model. Research Memorandum 76-16, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA, July 1976b.

Wheaton, G., Rose, A., Fingerman, P., Leonard, R., and Boycan, G. Evaluation of the effectiveness of training devices: Validation of the predictive model. ARI Technical Report TR-76-A2, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA, October 1976c.

Willis, M. and Peterson, R. Deriving training device implications from learning theory principles. Volume I: Guidelines for training device design, development and use. Technical Report NAVTRADEVCEN 784-1, U.S. Naval Training Device Center, Port Washington, NY, July 1961.